REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	9 September 1998	

4. TITLE AND SUBTITLE
REDUCTION OF INVALID ASSESSMENTS DURING CYCLE ERGOMETRY TESTING

5. FUNDING NUMBERS

6. AUTHOR(S)
WAYNE MARTIN PRITT

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF TEXAS

8. PERFORMING ORGANIZATION REPORT NUMBER

98-060

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With 35-205/AFIT Sup 1

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Approved for public release
Distribution Unlimited

DTIC QUALITY INSPECTED 1

14. SUBJECT TERMS

15. NUMBER OF PAGES
34

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

REDUCTION OF INVALID ASSESSMENTS DURING

CYCLE ERGOMETRY TESTING

By

WAYNE MARTIN PRITT, BA, MD

APPROVED:

_____

Jimmy L. Perkins, PHD

_____

Benjamin S. Bradshaw, PHD

DEDICATION

To God who has granted the talent _spacing_
and ability to succeed, not merely in this
endeavor, but in the rest of my life;

To my wife Brooke and my boys
Seth, Tyler and Chase who inspire love
and laughter despite trial and misfortune;

To my father and my mother, whose
accomplishment in successfully raising
my siblings and me in these amoral times
is a testament to duty, strength and love.

REDUCTION OF INVALID ASSESSMENTS DURING

CYCLE ERGOMETRY TESTING

By

WAYNE MARTIN PRITT, BA, MD

THESIS

Presented to the Faculty of the University of Texas

Health Science Center at Houston

School of Public Health

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF PUBLIC HEALTH

ACKNOWLEDGMENTS

I offer thanks to my advisors, Dr. Jimmy Perkins and Dr. Benjamin Bradshaw. They have kept this rickety train on track despite my best efforts toward derailment. I also gratefully acknowledge the support provided by the USAF Office for Prevention and Health Services (OPHSA) at Brooks Air Force Base, San Antonio, Texas without which this particular project would not be. Specifically, I thank Dr. Stefan Constable and Mr. Pete Flatten who provided considerable support and the precious data needed to make this research possible. Appreciation is also extended to my fellow classmates who have never failed to express concern and a willingness to help. Finally, and most importantly, I thank my wife who grudgingly granted the approval necessary for my return to academia at this point in our lives, and who unfailingly provided the "personal guidance" required when I attempted to take myself much too seriously.

Thesis submitted to the MPH Committee on the 30 June 1998.

REDUCTION OF INVALID ASSESSMENTS DURING

CYCLE ERGOMETRY TESTING

Wayne M. Pritt, BA, BA, MD, MPH
The University of Texas
Health Science Center at Houston
School of Public Health , *1998* .
~~San Antonio Campus, 1998.~~ —

Supervising Professor :   Jimmy L. Perkins

The cycle ergometry test (CET) was adopted by the United States Air Force in 1992 to provide an accurate, safe and practical means by which to screen the active duty population annually for aerobic fitness. Refinement of the test has ensued over the last six years, and perhaps the area of most active research is an attempt to minimize the number of CET evaluations that result in an inability to predict fitness, i.e. an invalid assessment. The magnitude of the problem with invalid assessments is revealed with the realization that historically up to 16.4 percent received an invalid result. Multiple causes exist for invalid results, each with a unique set of circumstances requiring correction.

This investigation focuses on the third most common cause for an invalid cycle ergometry assessment, specifically nonattainment of steady state heart rate. A review of an existing data set was undertaken using a proposed alternative definition for acceptable heart rate variation during the final two minutes of the test. By broadening the definition of steady state heart rate from $\pm$ 3 beats per minute to $\pm$ 5 or 6 beats per minute the invalid rate is dropped from 16.4 percent of all assessments to 2.48 or 1.06 percent respectively. This proposed change significantly decreases the invalid rate and fails to adversely affect the overall accuracy of the CET. Fewer required repeat evaluations will create a cost savings to the Air Force and help to generate increased confidence in the cycle ergometry test. In view of these benefits, this change should be adopted Air Force wide.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF APPENDICES

INTRODUCTION

Physical fitness, or the lack of it, has wide-ranging health ramifications for the individual and the public at large. We need look no further than the lay press to see a massing of information describing the therapeutic and prophylactic benefits of regular exercise. Accumulating scientific data support the notion that exercise and physical activity are related to the prevention and successful treatment of several chronic diseases (Thompson, 1994). The health of its individual members therefore clearly impacts the ability of the U.S. Air Force to prosecute its mission.

Active duty members of the U.S. Air Force (USAF) are mandated to defend the United States against all enemies. In order to be prepared to carry out this singularly important mission, the Air Force and its individual members must maintain a constant state of readiness. This state of readiness must include equipment and supplies, training and proficiency at required tasks, and the focus of this study, individual physical fitness. Lt. Gen. Charles Roadman II, Surgeon General of the United States Air Force, stated on Oct. 28 1997, that "Cycle ergometry is also a war-readiness issue". It is towards this end that approximately 350,000 USAF members are assessed annually by a cycle ergometry test (CET) to provide a screening assessment of their level of personal aerobic physical fitness.

PURPOSE OF THE STUDY

This investigation will measure the magnitude of the reduction in cycle ergometry invalid test results, manifested by a change in the definition of steady-state heart rate from

1

3 beats per minute to 5 or 6 beats per minute.   It will further assess the possible impact of such a change on the ability of the CET to accurately predict aerobic capacity.

BACKGROUND

The ability to do endurance-type muscular work is dependent on aerobic metabolism, i.e. the utilization of oxygen, customarily symbolized as VO_2. The maximum aerobic capacity is the maximum rate of oxygen utilization (VO_{2max}), and is the single best surrogate measurement for aerobic fitness. Determinants of oxygen utilization include the size of the muscle mass, the oxygen containing blood flow to the muscle, and the efficiency of oxygen extraction by the muscle. Cardiac output (CO) is defined as the amount of blood pumped by the heart per minute, and is determined by heart rate (HR) and stroke volume (SV) in the equation: $CO = HR \times SV$. VO_2 can then be defined by the equation: $VO_2 = CO \times (aO_2 - vO_2)$ where $aO_2\text{-}vO_2$ is the difference in oxygen content of the arterial blood entering a muscle mass and the venous blood returning to the heart. Endurance or aerobic exercise has been shown to cause increases in the SV and in the $aO_2\text{-}vO_2$ difference. VO_2max occurs when the cardiovascular system is maximally burdened and the muscles metabolic capacity fully engaged (i.e. CO and $aO_2\text{-}vO_2$ reach their highest levels). This is functionally measured by increasing the workload until no further work can be done and exhaustion occurs.

In the United States the "gold" standard for measurement of VO_2max is the maximal treadmill test. In other countries, where cycling is more popular, the maximal cycling test is preferred (Pollock, 1994). Actual measurement of VO_2max by either of

2

these two methods is time consuming. It requires considerable technical expertise and expensive laboratory equipment and procedures. In addition, a small but definite risk exists for the individual during the test. This risk has been estimated at one fatality for every 25,000 tests (Pollock, 1994). For these reasons, it is impractical to use direct measurement of VO_2max as a screening test for physical fitness.

The requirement for a simple, inexpensive, and safe screening test for aerobic fitness has been identified since the 1950's. A principle physiologic adaptation to aerobic training is that a given individual will have a lower heart rate response to exercise at a constant workload. Astrand and Rhyming utilized this principle in the development of a nomogram in 1954 that estimated VO_2max for a submaximal output while exercising on a cycle ergometer, a treadmill, or during a step test, by measuring heart rate and body weight (Astrand, 1954). The Astrand-Rhyming (AR) nomogram was later modified for use only on a cycle ergometer, and age corrections were incorporated to improve its accuracy for older individuals (Astrand, 1960). Multiple studies, yielding different modifications and improvements, have been performed with the original AR nomogram. Correlation coefficients, found during comparison of the AR nomogram with its various modifications and maximal treadmill testing, have ranged from $r = 0.39$ to 0.94 (Pollock, 1994).

Cooper (1968) developed a screening physical fitness test in the late 1960's employing a 12-minute run. It was later modified to a 1.5 mile run for ease of administration and adopted by the USAF for fitness screening in 1970. This test yielded

3

a correlation coefficient of r = 0.90 with maximal treadmill testing and was successfully employed by the USAF until 1992. In the mid 1980's concern developed about the adverse medical risk to active duty members taking the 1.5 mile run. Harrison (1980) states that the 1.5 mile run required levels of exertion equal to 60 percent of maximal effort and that this level incurred medical risk. Also of concern was the tendency for unfit individuals to force themselves beyond levels of safe exertion in order to obtain a passing time on the run. Anecdotal reports began to surface about fatalities associated with the testing method (Weber, 1989). In response, Sharp (1991) described a proposed medical prescreening program to avoid morbidity and mortality during testing that would have added significant cost in funds and time to the testing scenario. Due to these draw backs, the USAF adopted a modified AR submaximal cycle ergometry test (CET) in 1992 to replace the 1.5 mile run as a physical fitness screening test (Byrne, 1993).

DeWolfe (1997) describes the USAF CET in its present form. Initial workload is based on age, gender, weight, stated activity level (active = strenuous physical activity at least two times per week), and whether the individual smokes. A computer program developed at Armstrong Laboratory, Brooks AFB, aids in the standardization of the test (Hartung, 1993). Individual variables are entered describing gender, age, weight, and height, prior to the test. The individual is asked to pedal in cadence with a metronome at the frequency of 50 revolutions per minute. A two minute warm-up period precedes zero to three minutes of workload progression, during which the workload is increased in increments of 0.5 to 1.0 kp (kiloponds -unit of angular resistance, converted to units of

work when multiplied by revolutions per time) to achieve a steady-state heart rate. This final heart rate must by definition be greater than 126 beats/min, yet fall below 85 percent of the individuals predicted maximum heart rate. The predicted maximum heart rate is defined as the numerical result of subtracting their age in years from 220. The assessment is immediately terminated should the heart rate exceed this maximum.

Exercise is continued over the next six to seven minutes at this workload. Heart rate at the end of the test (HR_2) is compared with the heart rate from the end of the previous minute (HR_1). If HR_2 and HR_1 differ by more than three beats per minute, the evaluation is extended for an additional minute. Heart rate is again measured (HR_3) and compared to the heart rate from the previous minute (HR_2). A test is considered invalid, due to nonattainment of steady state, if the heart rate at the end of the assessment differs by more three beats from that recorded for the previous minute. A mean value is calculated for the steady state heart rate ($[HR_3-HR_2]/2$ or $[HR_2-HR_1]/2$) and recorded along with the steady state workload. The computer software then calculates the predicted VO_2max using an equation that describes the linear relationship between heart rate and VO_2max at a given workload. This equation is provided in appendix A.

Results of the CET were originally reported to participants by category, I–VI, or as Invalid. A passing score was defined as category III-VI. Category placement is defined by gender, age and VO_2max as estimated by the CET and displayed in Table 1. Reporting of results has since been changed. The six categories have been retained for data collection purposes, but participants are only told that they either pass or fail, or that their test was invalid (Byrne, 1993). The definition of a passing score is still retained at the level of

Table 1. USAF Cycle Ergometry Fitness Categories
Aerobic Capacity (VO_2) by Age for Men and Women

Fitness Category	Age (Men)			
	<29	30-39	40-49	>50
Category I	< 28.0	< 27.0	< 25.0	< 22.0
Category II	28.0 - 33.9	27.0 - 31.9	25.0 - 29.5	22.0 - 27.5
Category III	34.0 - 41.9	32.0 - 38.9	29.6 - 35.5	27.6 - 31.5
Category IV	42.0 - 47.9	39.0 - 45.9	35.6 - 41.5	31.6 - 36.5
Category V	48.0 - 54.9	46.0 - 52.9	41.6 - 47.5	36.6 - 42.5
Category VI	> 54.9	> 52.9	> 47.5	> 42.5
Fitness Category	Age (Women)			
	<29	30-39	40-49	>50
Category I	< 26.0	< 24.0	< 23.0	< 20.0
Category II	26.0 - 26.9	24.0 - 25.9	23.0 - 25.9	20.0 - 22.9
Category III	27.0 - 35.9	26.0 - 33.9	26.0 - 30.9	23.0 - 25.9
Category IV	36.0 - 42.9	34.0 - 38.9	31.0 - 36.9	26.0 - 30.9
Category V	43.0 - 48.9	39.0 - 46.9	37.0 - 40.7	31.0 - 34.9
Category VI	> 48.9	> 46.9	> 40.7	> 34.9

category III or better. Individuals successfully passing the CET are reevaluated annually. Those participants who fail are enrolled in remedial aerobic conditioning programs and retested upon completion of the program. Those individuals whose tests are scored as invalid are also retested, but without intervening aerobic training.

The proportion of individuals reported as passing the CET varies somewhat from study to study. Byrne (1993) quotes a 75 percent pass rate, but makes no mention of invalid tests. Lombard (1995) found that 42 percent of women and 64 percent of men pass on initial attempt, again with no mention of invalid results. The largest study to date (n = 9,437), performed by DeWolfe, (1997) found 74 percent passing and 9.6 percent failing. He further reports that 16.4 percent of participants were found to have an invalid test. Invalid test results were originally subdivided into seven possible outcomes as defined by the Air Force at that time, and described by DeWolfe (1997). He outlines the breakdown of invalid tests as follows, with the percentage of each outcome category as a function of the total number of invalid tests provided in brackets.

1) Heart rate exceeds 85% of maximum (max = 220 – age)   [39.7%]

2) Heart rate fails to reach 126 bpm in the last minute of the test   [7.8%]

3) Heart rate varies more than 3 bpm in the final two minutes   [15.1%]

4) Individual could not maintain 50 rpm on the cycle   [<1%]

5) Self-reported perceived exertion exceeds a score of 15   [8.9%]

6) Individual requests termination of the test   [1.6%]

7) Other   [26.5%]

The present requirement to repeat the cycle ergometry assessment for 16.4 percent of approximately 350,000 annual screenings (57,000), is an undesirable drain on personnel resources, time and funds. The relatively high proportion of invalid test results erodes the Air Force members confidence in the test and adversely affects morale (DeWolfe, 1997). Attempting to decrease the number of invalid tests is at present the focus of active interest.

The classification of invalid outcomes presented has since been modified to eliminate "perceived exertion" as a criteria for an invalid test. Participants may still request termination of the CET, but elimination of the implied organizational acceptance of excessive exertion as an acceptable reason for an invalid test should decrease the potential for misuse or abuse of the system.

Invalid outcome 1 is responsible for the largest percentage of invalid tests and is also being reviewed. Many individuals who exceed 85 percent of their maximum heart rate do so at a workload that would not have allowed them to pass even if they had not broken their heart rate ceiling. It would seem logical to reclassify their results as failures rather than invalids, since performance would not likely be improved with retesting in the absence of increased aerobic fitness. Those individuals who exceed their heart rate ceiling at workloads that otherwise would have resulted in a passing score, are the subject of experimentation on the timing and amount of workload progression initiated during the CET (DeWolfe, 1997).

Invalid outcome 7 constitutes the second largest proportion of invalid tests. These results are due to error by the test administrator or less frequently, equipment failure (computer hardware/software, or heart rate monitor). DeWolfe (1997) points out that the

newer, more automated software programs have a lower rate of outcome 7 invalid tests. Efforts underway toward more complete standardization in test administrator training and utilization of the improved software package, should ensure invalid tests due to outcome 7 are kept as low as possible.

As seen above, 15.1 percent of the tests scored as invalid occurred as a direct result of the definition of steady-state heart rate (HR varies by greater than 3 bpm in final two minutes of the test). A literature review indicates that the use of a 3 bpm "window" is not scientifically based. Indeed, other cycle ergometry tests using similar submaximal evaluations use 5 bpm as the cutoff for determination that a steady-state heart rate has been achieved (Pollock, 1994; Wisen, 1995). Redefining the acceptable steady-state heart rate as having been achieved when the heart rates recorded at the end of the last two minutes of the test are within 5 bpm, should reduce the number of invalid tests due to nonattainment of steady-state heart rate. Further, it should not be expected to adversely affect the validity of the overall test results, due to the daily personal variation in VO_2max (Pollock, 1994).

METHODS AND PROCEDURES

A data set was obtained from the USAF Office for Prevention and Health Services (OPHSA) at Brooks Air Force Base, San Antonio, Texas. This data set is the same as utilized in the analysis of overall invalid CET results as published by DeWolfe (1997). The data consists of 9,496 cycle ergometry test assessments from five Air Force Bases: Brooks AFB, Kelly AFB, Lackland AFB, and Randolph AFB in Texas, and Patrick AFB in Florida obtained during the 1996 fiscal year.

9

Personal identifiers were associated with the raw data and it was therefore maintained under lock and key to protect individual privacy for the duration of the present investigation. Data were returned to OPHSA at the completion of the present study. As this was an extant data set, individual consent was neither pursued nor obtained.

All active duty U.S. Air Force members are required to undergo annual physical fitness assessment as evaluated by the cycle ergometry test. Eligibility for selection in the original study, and therefore by default this study, was based on duty assignment to one of the five previously delineated bases, and completion of the requirement for routine annual cycle ergometry assessment. The data therefore represents a cross-section of the active duty U.S. Air Force population at these five bases to include officer and enlisted, ages 18 to 67, and male and female participants. In so far as the individuals assigned to the bases used to collect data are representative of the U.S. Air Force in total, the study participants represent the active duty population at large. The data were first analyzed to describe the demographics of the population studied as it pertains to gender, age, and rank. The demographic characteristics of the study participants were then compared to the total Air Force in order to argue external validity.

Data were then analyzed to obtain the overall rate of all invalid tests among all CET assessments and the specific invalid rates by outcome category among all invalid tests. This served to validate the previous findings of DeWolfe and to allow examination of the magnitude of the problem with invalid CET assessments in general and outcome category 3 specifically.

Further analysis focused on the invalid assessments in outcome category 3. All individual test results within this category were reassessed for the presence of acceptable steady-state heart rate by substituting the proposed study definition of less than or equal to 5 beat per minute variation within the last two minutes of the test for the presently used 3 beat per minute standard. These same invalid CET assessments were then reviewed using a second steady-state heart rate definition of less than or equal to 6 beats per minute.

Additional analysis sought to elucidate the observed frequency of the variation in the heart rate recorded at the termination of the evaluation and one or two minutes prior to the end of the test. These results were used to graphically evaluate the distribution heart rate variation.

RESULTS

Demographic Characteristics

Demographic comparison of the study group and the U.S. Air Force Population as a whole are displayed in Table 2. Review of these data reveals that the study group is slightly older and held slightly higher rank than the total Air Force population. The largest variation appears in gender, with the study group containing 5 percent more females than the total Air Force. None the less, the study group appears to reflect the population of the Air Force adequately. Conclusions drawn from the analysis should therefore be generalizable to the Air Force as a whole.

11

**Table 2. Demographic Comparisons of Gender, Rank and Age
Between the Study Group from Five Air Force Bases
and the Total Air Force Population**

	Study Group		Air Force	
	#	%	#	%
Gender				
Male	7,327	**77.16%**	302,253	**82.27%**
Female	2,169	**22.84%**	65,147	**17.73%**
Rank				
E1-E5	4,785	**50.81%**	218,437	**59.46%**
E6-E9	2,149	**22.82%**	76,886	**20.93%**
O1-O3	1,558	**16.54%**	42,110	**11.46%**
O4-O7	925	**9.82%**	29,967	**8.16%**
Other	79		N/A	
Age				
17 - 19	101	**1.06%**	19,262	**5.24%**
20 - 24	1,775	**18.69%**	88,238	**24.02%**
25 - 29	2,052	**21.61%**	75,501	**20.55%**
30 -34	2,060	**21.69%**	71,412	**19.44%**
35 - 39	1,984	**20.89%**	70,606	**19.22%**
40 - 44	1,049	**11.05%**	30,616	**8.33%**
45 - 49	359	**3.78%**	9,420	**2.56%**
50 +	116	**1.22%**	2,344	**0.64%**

Invalid Assessments

The overall percentage of invalid assessments due to all categories was found to be 1740 of 9,496 tests, or 18.32 percent. The specific rates of invalid cycle ergometry assessments by outcome category are displayed in Figure 1. Failure to obtain a steady state heart rate, outcome category 3, is the third most frequent cause of an invalid CET assessment at 13.97 percent, and found to be 2.56 percent of all assessments performed. Only invalid assessments in outcome categories 1 (38.33 percent) and 7 (30.52 percent) exceed this proportion. These data vary somewhat from that published by DeWolfe (1997). The variation is explained by the fact that while this investigation reviewed all CET results obtained, regardless of the number of times an individual attempted the test, he did not include repeat assessments from the same individual.

Outcome Category 3 Invalid Assessments

Examination of the data set reveals the frequency distribution of heart rate variation in excess of 3 beats per minute as displayed in Figure 2. It can be seen that 83.97 percent of the variations exceeding 3 beats per minute are less than or equal to 5 beats per minute. Similarly, using less than or equal to 6 beats per minute as the standard for acceptable heart rate variation would capture 93.25 percent of presently invalid CET assessments.

Instituting a change in the definition of acceptable steady state heart rate variation results in a decrease in the number of outcome category 3 assessments as shown in Table 3. Using 5 beats per minute as the accepted standard results in a decrease in the number of invalid assessments from 243 to 38, an 84 percent reduction. Likewise, utilizing a 6 beat per minute standard yields a 93 percent reduction (243 to 16) in the number of invalid

13

Figure 1. Distribution of Invalid Cycle Ergometry Testing
Assessments in the USAF by Outcome Category

Outcome 1
Outcome 2
Outcome 3
Outcome 4
Outcome 5
Outcome 6
Outcome 7

38.33%

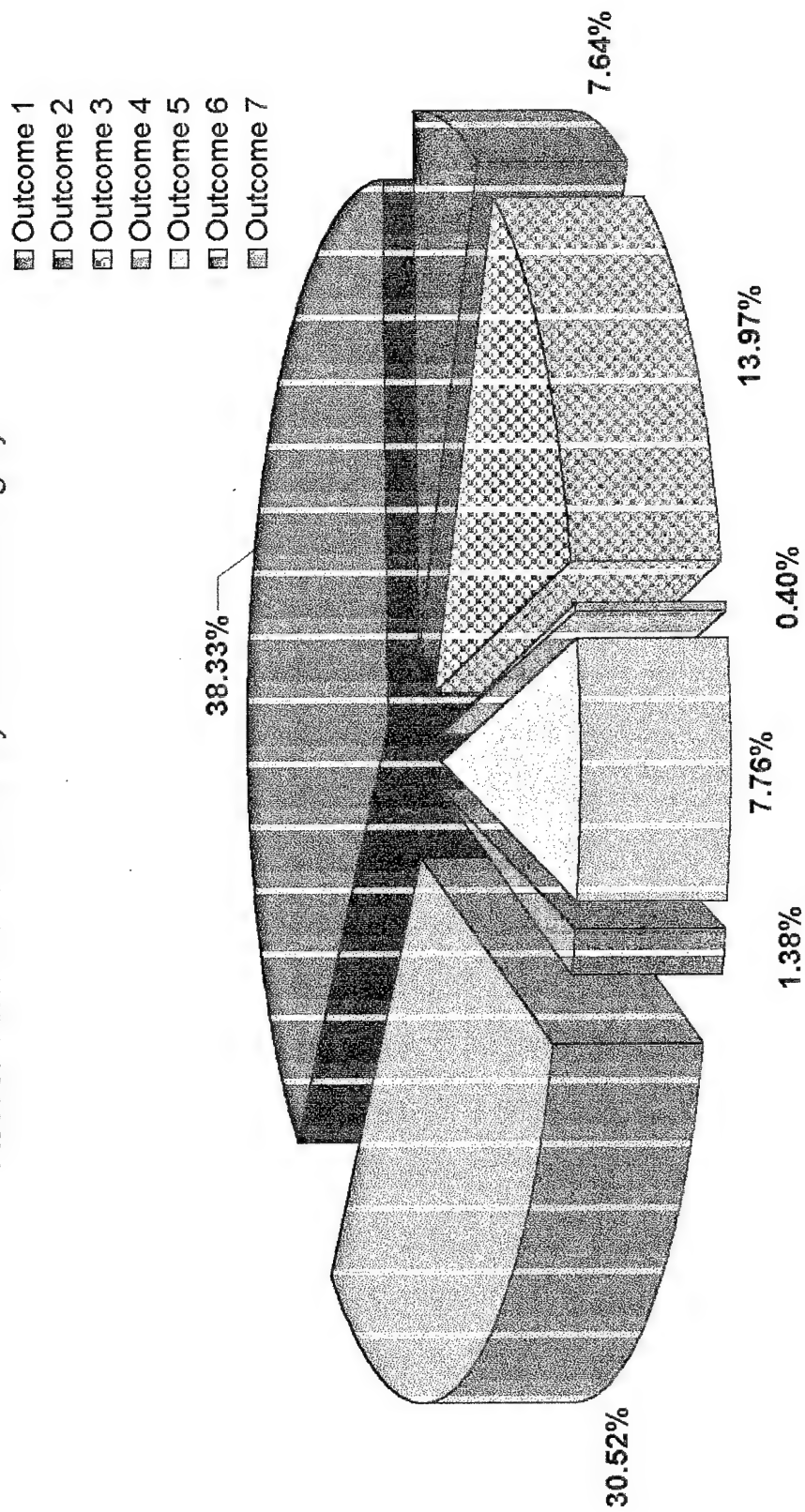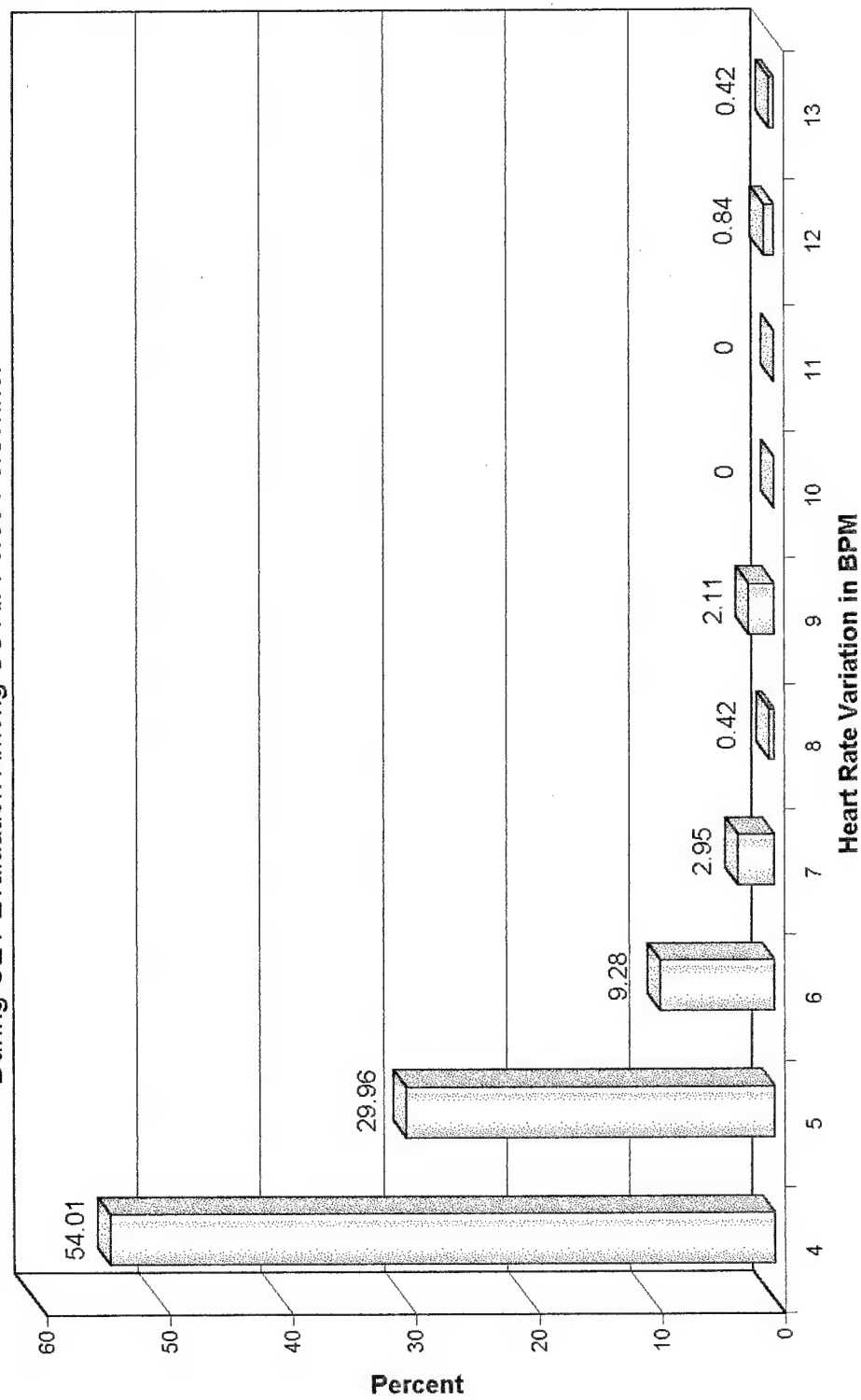13.97%

7.64%

0.40%

7.76%

1.38%

30.52%

Figure 2. Percentage Frequency of Heart Rate Variations > 3 BPM During CET Evaluation Among US Air Force Personnel

assessments due to nonattainment of steady state heart rate. This reduces the contribution

to the total number of CET invalid assessments by outcome category 3 from 13.97 percent

to 2.48 percent and 1.06 percent respectively.

DISCUSSION OF METHODS AND PROCEDURES AND RESULTS

Demographic Characteristics

The discrepancy between the study group and the Air Force as to age, rank and

gender is not overwhelming and should not impact the external validity of this

investigation. The difference may however, be attributed to two possible causes. First,

detailed demographic information describing the Air Force in 1996 is not available. The

demographic information for the U.S. Air Force describes the force as of March 1998,

while data for the study were collected during 1996.. This obviously impacts the

comparability of the two populations in regard to all three descriptors. Second, reductions

in the size of the active duty force have occurred in the intervening two years. These

reductions have predominately occurred through involuntary separations and early

retirements, with only small changes in the number of new Air Force members. These

changes as described would have disproportionately decreased individuals with longer

time in service and therefore of older age and more senior rank. The net effect would be

that the study group would appear to be older and of more senior rank than the U.S. Air

Force as a whole, just what the demographic analysis demonstrates. In summary, the

magnitude of the differences seen between the two populations is small, may be

16

O no # ↓rnt out of sequence

Table 3. Comparison of the Rates of Invalid Cycle Ergometry Testing Using Two Alternative Standards for Allowed Steady State Heart Rate Variation and the Present 3 BPM Standard

Allowed HR Variation	# Invalid	% of All Tests	% of All Invalids
≤ 3 bpm	243	2.56%	13.97%
≤ 5 bpm	38	0.40%	2.48%
≤ 6 bpm	16	0.17%	1.06%

Total Test Results Analized = 9,496
Total Invalids All Categories = 1740

17
~~18~~

explainable, and should not negatively affect the ability to generalize conclusions about the study group to the total Air Force population.

Outcome Category 3 Invalid Assessments

A more liberal definition of the acceptable variation in steady-state heart rate should logically decrease the rate of invalid cycle ergometry assessments due to nonattainment of steady-state heart rate, i.e. outcome category 3 invalids. The question before us is by how much, and will it affect the validity of the CET?

The results presented clearly elucidate the probable magnitude of the benefit of broadening the definition of steady state heart rate. Redefining acceptable variation in the steady state heart rate as plus or minus five or six beats per minute, will decrease the percentage of invalid cycle ergometry tests due to outcome category 3 from 2.56 percent of all CET assessments to 0.40 or 0.17 percent respectively. Given that as of March 1998, the Air Force conducts 367,400 tests annually, this equates to a reduction in the number of invalid tests due to this problem from 9,405 at the present rate to 1,470 ($\pm$ five bpm standard) or 625 ($\pm$ six bpm standard) annually. This means that at least (many repeat assessments are themselves invalid for the same or other reasons) 7713 to 8565 repeat CET assessments could be avoided by adopting this proposed change to the testing protocol.

The next logical question is whether making this change in the testing protocol affects the validity of the cycle ergometry assessment and the estimation of VO_2max it ultimately provides. It was previously stated that making the change to five beats per minute should not affect the overall validity of the cycle ergometry test (Pollock, 1994). A

logical argument follows to justify this statement and to argue possible extension to 6 beats per minute.

A basic physiologic principle is that an individual with an aerobically conditioned cardiovascular system will exhibit a lower heart rate at a given workload than an unfit individual. Further, the assumption is made that the heart rate increases linearly with increasing workload to the point of maximum aerobic capacity. A steady state heart rate is therefore utilized to determine when the cardio-pulmonary system has reached equilibrium in response to a given workload. This steady state heart rate is then correlated with the known workload and the aerobic capacity, or VO_2max, is estimated using equations describing this linear relationship between heart rate and workload. Unfortunately, not all individuals equilibrate rapidly, or at the same rate.

It is theoretically possible to extend the testing protocol in order to allow more time for cardio-pulmonary equilibrium to be reached. Muscular fatigue may, however, intervene, thereby disrupting the useful extension of the test and the possibility of obtaining a measurement of VO_2max. Problems may also arise with the standardization of the assessment when the length of the test is open-ended. As a result, the length of the assessment is restricted, and the measurement of steady state heart rate, in the midst of exercise, is estimated as the mean value between the final heart rate, and the heart rate recorded two minutes prior to the completion of the test. This assumes, of course, that the difference in the two heart rates does not vary by more than the *a priori* determined amount, i.e. not more than three beats per minute in the present U.S. Air Force protocol.

Expanding the steady state heart rate zone to five or six beats per minute would have only a small impact on the determination of predicted VO_2max. A series of three hypothetical scenarios will help to illustrate this point. Assume for all three scenarios that a 30 year old male, weighing 170 pounds and standing five feet ten inches tall is undergoing cycle ergometry evaluation, and has reached a steady-state heart rate at a 3.0 kp work load. In the first scenario his heart rate on completion of the test (HR_2) and his heart rate the previous minute (HR_1) does not vary and is 140 beats per minute. The equations describing the linear relationship of heart rate and aerobic capacity reveal his VO_2max to be predicted as 48.7 ml/kg/min. In the second scenario, now assume he does not reach perfect cardio-pulmonary equilibrium, but rather his heart rates are 140 bpm and 143 bpm respectively, with a mean value of 141.5 bpm. The nomogram equation now predicts a VO_2max value of 48.0 ml/kg/min. Finally, assume this time his heart rate the minute prior to the end of the test is 140 bpm, and the heart rate in the final minute of the assessment is 145 bpm. His mean steady-state heart rate value is now calculated to be 142.5 bpm, with his VO_2max predicted by the equation as 47.5 ml/kg/min.

Comparison of the predicted VO_2max estimations from these three scenarios reveals a difference in the first and the second values of 0.7 ml/kg/min (1.44 percent), and a difference in the first and the third scenarios of 1.2 ml/kg/min (2.46 percent). Differences of such magnitude are well within the known range of 1 to 3 ml/kg/min (2 to 5 percent) daily variation of predicted VO_2max estimations for an individual (Pollock, 1994). In other words, the degree of imprecision introduced in the VO_2max prediction by the suggested change is less than the total innate error of the test estimation. Therefore,

changing the definition of steady-state heart rate from $\pm$ three bpm to $\pm$ five bpm can be seen to have minimal impact of the overall accuracy of the cycle ergometry test and it's ability to predict maximum aerobic capacity. Figure 3 provides a graphical representation of this argument extended over a range of possible heart rates using the present three beat per minute standard, and the proposed five or six beat per minute alternatives. Table 4 displays the corresponding magnitude of the percentage changes in VO_2max prediction over the same range of heart rates. Finally, Figure 4 illustrates the percentage change in predicted VO_2max for a range of heart rates using three, five and six beat per minute allowed variations in steady state heart rate.

It is interesting to note that the magnitude of the percentage change in VO_2max predictions (as depicted in Figure 4) decreases with increasing heart rate both within the same standard and between the three standards. This implies that the maximum error in VO_2max prediction due to variation in steady state heart rate occurs at the lowest heart rates. This realization is important for the following reasons. First, the most aerobically fit individuals have the lowest heart rates at a given workload. Second, as a screening test, the CET is less interested in providing an exact measure of predicted aerobic fitness than it is in identifying fit from non-fit. Third, given that more accuracy (less error) is available at the higher heart rate ranges (the more un-fit individuals), these changes will have the least amount of negative impact on the less aerobically conditioned, maintaining the ability to discern between fit and non-fit individuals.

Also germane to the discussion concerning the acceptable limit for variation in steady state heart rate is the magnitude in the variation of heart rate inherent in the

Figure 3. Comparison of Predicted VO2max vs Steady State Heart Rate using 0, 3, 5 and 6 BPM Definitions of Acceptable Variation Between Final Heart Rate (HR2) and Heart Rate of Previous Minute (HR1) Amoung US Air Force Personnel During CET Evaluation

Steady State Heart Rate (bpm)
[(HR2-HR1)/2]

Predicted VO2max (ml/kg/min)

[Assumes gender=male, age=30yrs, weight=170lbs, height=70in, final workload=3.0kp]

Legend:
— HR2 - HR1 = 0 bpm
— HR2 - HR1 = 3 bpm
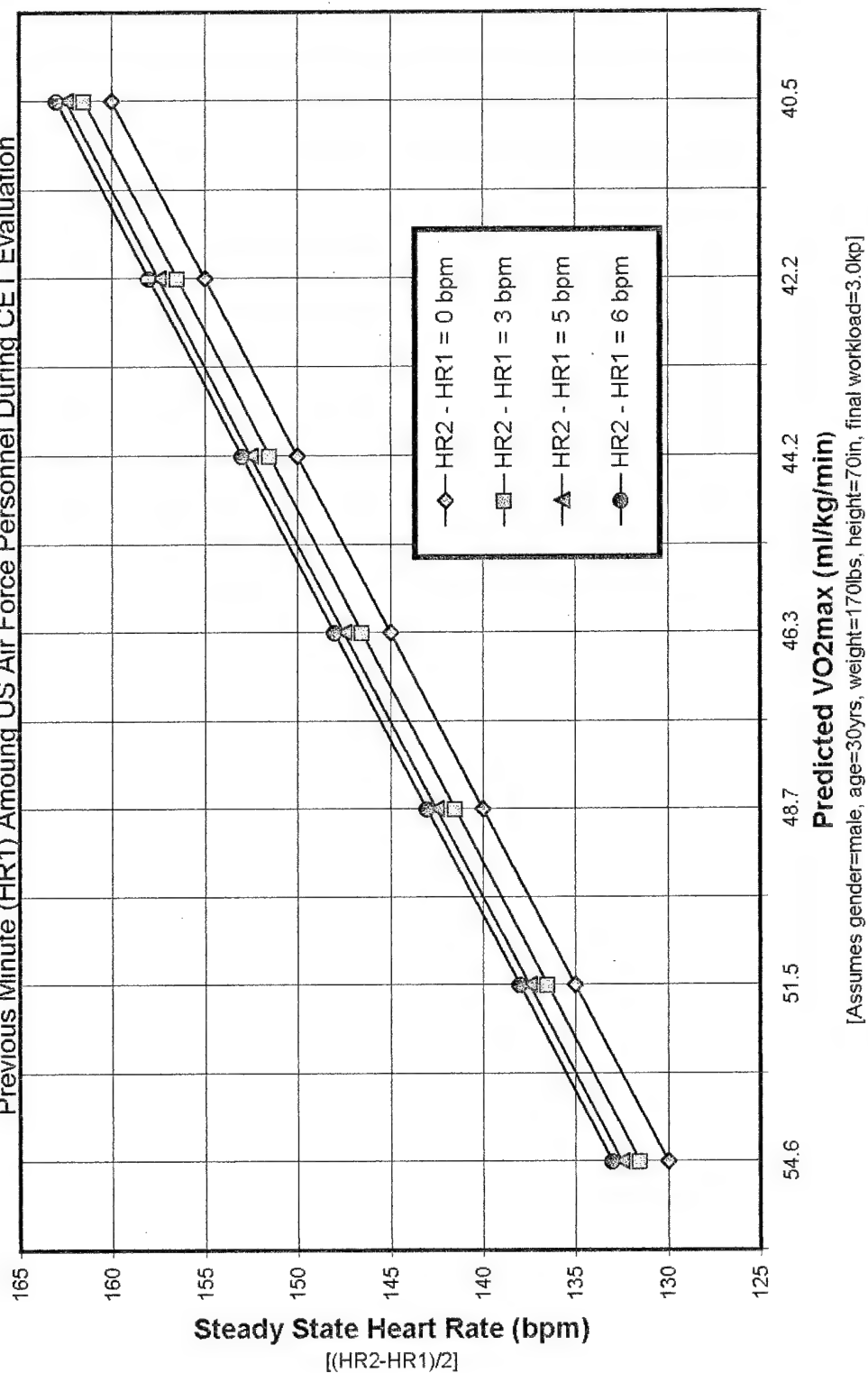— HR2 - HR1 = 5 bpm
— HR2 - HR1 = 6 bpm

Table 4. A Comparison of the Percentage Change in Predicted VO_2max Using 3, 5, and 6 BPM Standards for Acceptable Steady State Heart Rate Variation at Several Representative Heart Rates During Cycle Ergometry Testing in Air Force Personnel

"Heart Rate"[2] (bpm)	Percent Change in Predicted VO_2max[1]				
	0 vs 3 bpm	0 vs 5 bpm	0 vs 6 bpm	3 vs 5 bpm	3 vs 6 bpm
130	1.83	2.93	3.48	1.12	1.68
135	1.75	2.72	3.30	0.99	1.58
140	1.44	2.46	3.08	1.04	1.67
145	1.51	2.38	2.81	0.88	1.32
150	1.36	2.26	2.72	0.92	1.38
155	1.18	1.90	2.37	0.72	1.20
160	1.23	1.98	2.22	0.75	1.00

[1] Assumes gender=male, age=30yrs, weight=170lbs, height=70in, workload=3.0kp
[2] Actual heart rate (HR) used in calculations is (depicted "HR" + defined HR variation)/2

Figure 4. Comparison of the Changes in Predicted VO2max Incurred by Differing Standards of Acceptable Steady State Heart Rate Variation During CET Evaluation in the US Air Force
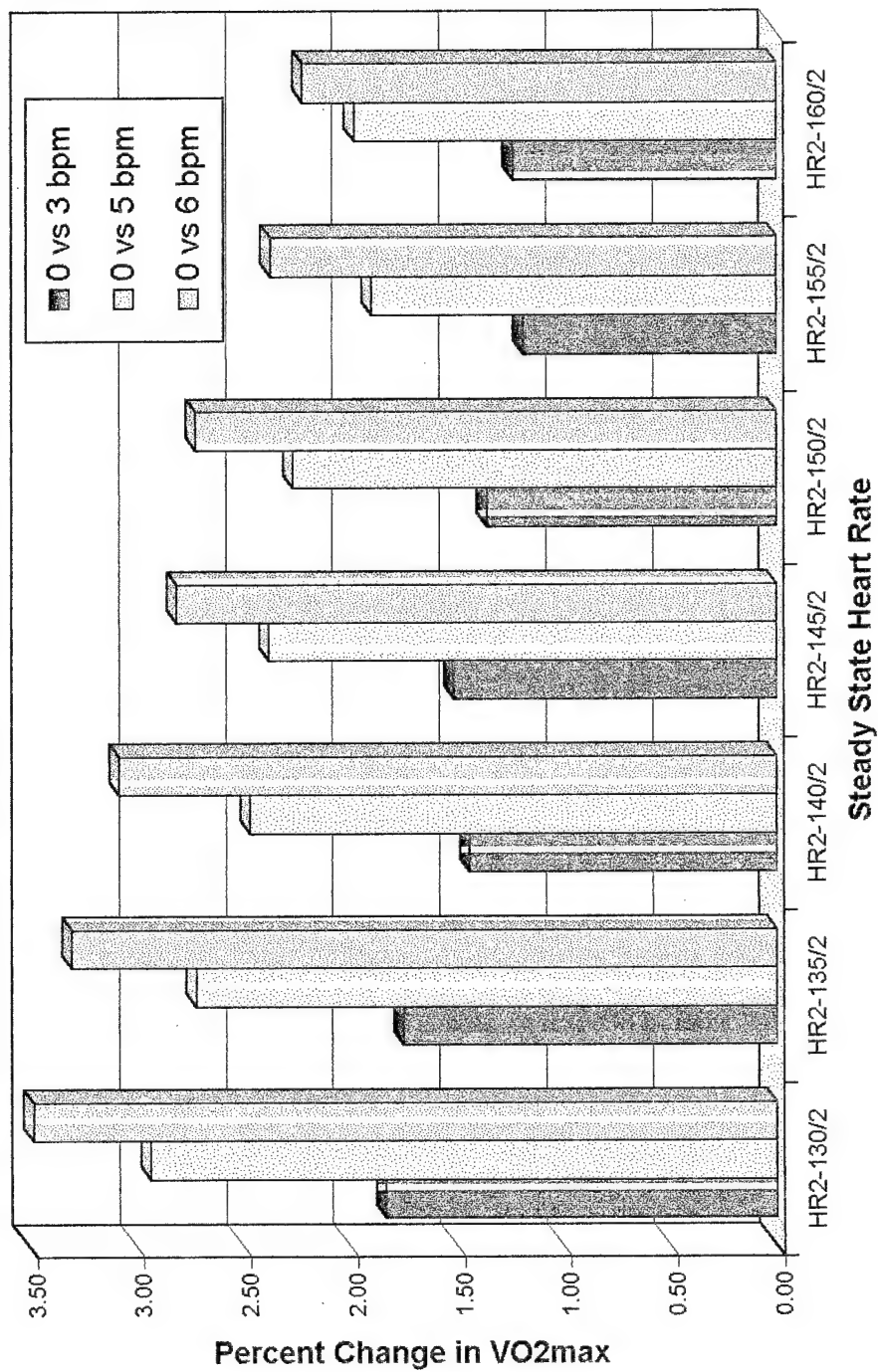
individual, i.e. the intra-individual variation in heart rate. If repeated measurements of heart rate were made of an individual at constant levels of exertion, it would be possible to obtain the standard deviation of the measurements around the mean. Using the standard deviation it is possible to calculate the range in which 95 percent of repeated measurements would fall as $\pm$ [2.77] x [standard deviation] (ASTM, 1990). This means the standard deviation of intra-individual heart rates need be no larger than 1.8 in order to justify using the five beat per minute standard, or no larger than 2.2 to support the six beat per minute standard. [$\pm$ 5 = (2.77) x (standard deviation)] Unfortunately, though the literature contains a plethora of information concerning beat to beat heart rate variations, these data describe populations and not individuals, and therefore can not be used in this calculation. Despite the lack of success in the literature search for this information, it is difficult to believe that these measurements have not been taken and recorded somewhere as a means to an end during other physiologic research.

Cost Savings

Fewer tests would be required annually to screen the active duty force, given an 84 to 93 percent decrease in outcome category 3 invalid assessments. It stands to reason that this would create some degree of monetary cost savings for the Air Force. This investigation revealed a rate of 2.56 percent for the overall occurrence of category 3 invalid assessments. Assuming the present personnel strength of the Air Force of 367,400, this means 9,405 assessments will be invalid due to excessive variation of steady state heart rate. If however, you apply the $\pm$ five beat per minute standard, 7,898 (83.97 percent) of these invalid assessments are now valid and do not have to be repeated, the $\pm$

six beat per minute standard saves 8,771 (93.25 percent). Assuming an estimated cost for each CET at $38.47, this equates to an annual savings to the Air Force of from $303,836.00 to $337,420.00. Appendix B provides a brief outline of the rationale used for these estimates.

CONCLUSIONS AND RECOMMENDATIONS

The U.S. Air Force has a legitimate need to encourage, maintain and periodically assess the physical fitness of its active duty members. Given the current Air Force philosophy of minimizing medical risk to the individual during fitness testing, the cycle ergometry test is a logical and sound choice of assessment. Problems with the CET, however, exist in the undesirably high rate of invalid tests. Invalid assessments are costly, time consuming, and adversely affect U. S. Air Force personnel morale and confidence in the test by generating the requirement for repeat evaluations. Invalid assessments are generated by a variety causes and each different cause will require a unique strategy to reduce its impact on the overall rate of invalid tests. A change in the definition of what constitutes an acceptable range for heart rate variation in the determination of steady state heart rate has the potential to reduce the number of invalid cycle ergometry tests by 84 to 93 percent. Further, adoption of the proposed changes would result in a savings of over $300,000.00 per year. This proposed modification of the current testing protocol would be easily accomplished and could be quickly implemented. It has bee shown to have minimal effect on the cycle ergometry tests ability to accurately predict an individuals level of aerobic fitness. In light of these facts, the U.S. Air Force should seriously consider

this proposed modification of the current protocol to partially achieve its goal of reduced

invalid cycle ergometry assessments.

Appendix A. Equations for Calculation of VO$_2$max

Definitions. The following definitions of terms apply for calculating VO$_2$max:

$a \equiv$ age of the individual in years
$C_{BSA} \equiv$ conversion constant
$h \equiv$ height of the individual in cm
$H_{X,avg} \equiv$ average heart rate of the individual in bpm
$H_X(n) \equiv$ the individual's heart rate in bpm at the end of minute n
$m \equiv$ weight of the individual in kg
$W_{L,end} \equiv$ ending workload in kp
$V_{O2,raw} \equiv$ raw VO$_2$ score in ml/kg-min

Calculation for females. The estimated VO$_2$max for females is calculated as follows:

$$X_A = \frac{3054 \cdot W_{L,end} + 670 \cdot C_{BSA}}{5} \text{ ml O}_2/\text{kg-min},$$

where

$$C_{BSA} = 0.004591 \cdot h^{3/4} \cdot m^{1/2}$$

$$X_B = \frac{210 - 72}{H_{X,avg} - 72} = \frac{138}{H_{X,avg} - 72}$$

$$X_C = \frac{100}{100 + (1.14 \cdot a) - 23}$$

$$X_D = -5 \cdot \left[\frac{\dfrac{X_A \cdot X_B \cdot X_C}{m} - LL}{UL - LL} \right] \text{ ml O}_2/\text{kg-min},$$

where the bracketed value is $0 \leq [\,] \leq 1$, and values exceeding the limits are rounded to the nearest limit value. The upper and lower limits, UL and LL respectively, are obtained from Table A.

Age	Females		Males	
(years)	Lower Limit	Upper Limit	Lower Limit	Upper Limit
17 - 24	26	49	28	55
25 - 29	26	48	28	54
30 - 34	25	47	27	53
35 - 39	24	44	27	51
40 - 44	23	41	26	48
45 - 49	22	38	24	46
50 - 54	21	35	23	43
55 - 59+	19	35	21	43

Table A. Estimation correction factor limits (ml O_2/kg-min).

$$VO_2 \max(female) = \frac{(X_A \cdot X_B \cdot X_C)}{m} + X_D \text{ ml } O_2/\text{kg-min}$$

Calculation for males. The estimated VO_2max for males is calculated as follows:

$$X_A = \frac{3054 \cdot W_{L,end} + 670 \cdot C_{BSA}}{5} \text{ ml } O_2/\text{kg-min},$$

where

$$C_{BSA} = 0.004591 \cdot h^{3/4} \cdot m^{1/2}$$

$$X_B = \frac{205 - 61}{H_{X,avg} - 61} = \frac{144}{H_{X,avg} - 61}$$

$$X_C = \frac{100}{100 + (1.37 \cdot a) - 33.2}$$

$$X_D = 5 \cdot \left[\frac{UL - \frac{X_A \cdot X_B \cdot X_C}{m}}{UL - LL} \right] \text{ ml } O_2/\text{kg-min},$$

where the bracketed value is $0 \leq [\ ] \leq 1$, and values exceeding the limits are rounded to the nearest limit value. The upper and lower limits, UL and LL respectively, are obtained from Table A.

$$VO_2 \max(male) = \frac{(X_A \cdot X_B \cdot X_C)}{m} + X_D \text{ ml } O_2/\text{kg-min}$$

Appendix B. CET Cost Estimate Calculations

Assume-
workday = 8 hours
time to perform 1 CET = 1 hour   → 8 tests/day

Cost of Test Administrator:
assume-
 E-4 is the average rank of test administrator
hourly wage E-4 = 15.62   → $15.62/test

Cost of Testee:
assume-
average rank officer = Capt
hourly wage Capt = 37.15
average enlisted rank = E-5
hourly wage E-5 = 18.81
4:1 ratio enlisted to officer in AF
average cost for testee   → 4(18.81) + 1(37.15) = $22.48/hr

Cost of Equipment:
> Bike
>> Lifetime 10 years, initial cost $850.00
>> 10 years = 2600 work days
>> 8 test/day * 2600 days = 20,800 tests
>>> → $850.00/20,800 tests = $0.04/test

> Heart Rate Monitor
>> Lifetime 3 years, initial cost $90.00
>> 3 years = 780 work days
>> 8 test/day * 780 days = 6240 tests
>>> → $90.00/6240 tests = $0.01/test

> Computer
>> Lifetime 3 years, initial cost $2000.00
>> 3 years = 780 work days
>> 8 test/day * 780 days = 6240 tests
>>> → $2000.00/6240 tests = $0.32/ test

Total Cost per Test:

$15.62 + $22.48 + $0.04 + $0.01 + $0.32 = **$38.47/test**

This estimate does not include overhead, space costs, or loss of efficiency in primary work place due to time away from the job (i.e. what is the cost of a pilot who can't fly because he has to take his CET again?)

BIBLIOGRAPHY

ASTM: "Standard Practice for Use of the Terms Precision and Bias in ASTM Test Methods", E 177 – 90a, 1990, pp. 152

Astrand, PO., Rhyming, I., "A Nomogram for Calculation of Aerobic Capacity (Physical Fitness) From Pulse Rate During Submaximal Work", Journal of Applied Physiology, Vol. 7, 1954, pp. 218-221

Astrand, I., "Aerobic Work Capacity in Men and Women with Special Reference to Âge", Acta Physiology Scandinavia, Vol. 49. Suppl. 169, 1960

Byrne,TJ., Block, PA., "How does Cycle Ergometry Measure Up?", USAF Medical Service Digest, Vol. 44, No. 4, 1993, pp. 8-10

Cooper, KH., "Testing and Developing Cardiovascular Fitness Within the United States Air Force", Journal of Occupational Medicine, Vol. 10, No.11, 1968, pp. 636-639

Cooper, KH., Zechner, A., "Physical Fitness in United States and Austrian Military Personnel, A Comparative Study", Journal of the American Medical Association, Vol. 215, No. 6, 1971, pp. 931-934

DeWolfe, G., Flatten, P., Chao, S., O'dowd, R., Richardson, D., Snedecor, M., Constable, S., "Invalid Cycle Ergometry Assessment Outcomes at Five Air Force Bases", USAF Armstrong Laboratory Document No. AL/PS-TR-1997-0056, April 1997

Fernhall, B., Kohrt, W., "The Effect of Training Specificity on Maximal and Submaximal Physiological Responses to Treadmill and Cycle Ergometry", The Journal of Sports Medicine and Physical Fitness, Vol. 30, No. 3, 1990, pp.268-275

Glassford, RG., Baycroft, GHY., Sedwick, AW., Macnab, RBJ., "Comparison of Maximal Oxygen Uptake Values Determined by Predicted and Actual Methods", Journal of Applied Physiology, Vol. 20, 1965, pp. 509-514

Harrison, MH., Bruce, DL., Brown, GA., Cochrane, LA., "A Comparison of Some Indirect Methods for Predicting Maximal Oxygen Uptake", Aviation, Space, and Environmental Medicine, Vol. 51, No. 10, 1980, pp. 1128-1132

Hartung,GH., Krock, LP., Crandall, CG., Bisson, RU., Myhre, LG., "Prediction of Maximal Oxygen Uptake from Submaximal Exercise Testing in Aerobically Fit and Nonfit Men", Aviation, Space and Environmental Medicine, Vol. 64, No. 8, 1993, pp. 735-740

Hermiston, RT., Faulkner, JA., "Prediction of Maximal Oxygen Uptake by a Stepwise Regression Technique", Journal of Applied Physiology, Vol. 30, No. 6, 1971, pp. 833-837

Jessup, GT., Tolson, H., Terry, JW., "Prediction of Maximal Oxygen Intake From Astrand-Rhyming Test, 12 Minute Run, and Anthropometric Variables Using Stepwise Multiple Regression", American Journal of Physical Medicine, Vol. 53. No. 4, 1974, pp. 200-207

Lombard, DN., Lombard, TN., Poston, WSC., "A Behavioral and Environmental Intervention to Reduce False-Positives in Submaximal Bicycle Ergometry Testing", Military Medicine, Vol. 160, No. 7, 1995, pp. 347-351

Pollock, ML., Garzarella, L., DeHoyos, D., Brechue, W., Beekley, M., Werber, G., Lowenthal, DT., "The Cross-Validation of the United States Air Force Submaximal Cycle Ergometry Test to Estimate Aerobic Capacity", USAF Armstrong Laboratory Document AL/CF-TR-1994-0046, June 1994

Schlesselman,J., "Sample Size Requirements in Cohort and Case Control Studies of Disease", American Journal of Epidemiology, Vol. 99, No. 6, 1974, pp. 381-384

Sharp, JR., "The New Air Force Fitness Test: A Field Trial Assessing Effectiveness and Safety", Military Medicine, Vol. 156, No. 4, 1991, pp. 181-185

Teraslinna, P., Ismail, AH., MacLeod, DF., "Nomogram by Astrand and Rhyming as a Predictor of Maximum Oxygen Intake", Journal of Applied Physiology, Vol. 21, 1966, pp. 513-515

Thompson,WG., "Exercise and Health: Fact or Hype?", Southern Medical Journal, Vol.78, No. 5, 1994, pp. 567-574

Warren, JH., Jaffe, RS., Wraa, CE., Stebbins, CL., "Effect of Autonomic Blockade on Power Spectrum of Heart Rate Variability During Exercise", American Journal of Physiology, Vol. 273, No. 2, 1997, pp. R495-R502

Weber, S., "Faster Run Rules Worry Top Doctors", <u>Air Force Times</u>, Vol. 9, No. 11, 1989, pp. 15

Williford, HN., Sport, K., Wang, N., Olson, MS., Blessing, D., "The Prediction of Fitness Levels of United States Air Force Officers: Validation of Cycle Ergometry", <u>Military Medicine</u>, Vol. 159, No. 3, 1994, pp. 175-178

Wisen, AGM., Wohlfart,B., "A Comparison Between Two Exercise Tests on Cycle; A Computerized Test Versus the Astrand Test", <u>Clinical Physiology</u>, Vol. 15, 1995, pp. 91-102